Chapter 7: Job Priority and Queue Management

A library manager receives requests for file read and write jobs from users via **encp**, stores these requests in a queue, assigns a priority to each job, and dispenses the jobs to the mover for actual data transfer. A job's priority determines when it will get processed relative to other jobs in the queue, and its priority may change as circumstances change while the job waits in the queue. There are four items that factor into determining priority: category, numerical value within the category, "Fair Share", and ownership of resources by a group/experiment.

7.1 Job Priority Categories

There are two categories of job priority: normal and DAQ/Admin. The default priority is "normal". DAQ/Admin priority, as its name implies, is reserved for high-priority jobs.

Jobs that come in with DAQ/Admin priority get moved to the head of the queue. If there are more than one at a time, the other priority-related factors determine the order in which they get processed. Any normal priority job that is in progress is allowed to complete, but the system does not then process other normal jobs in the queue that are waiting for the same volume. At the completion of the current job, the tape is replaced in the drive if necessary, then the (first) DAQ/Admin job gets processed.

DAQ/Admin priority is granted only to jobs that satisfy certain preconfigured conditions. Conditions, if set, filter on the job's originating username, group, node, and so on. For example, conditions could be set to grant DAQ/Admin priority to all jobs submitted by user *joe* from node *mynode1*, by users *george* or *kim* from any node, and by any user from node *myspecialnode*; all other jobs get normal priority.

Setting conditions for DAQ/Admin priority is an administrative function.

7.2 Numerical Priority Values

Within its priority category, each job is also assigned an initial numeric priority value. The numeric priority is set by default according to preconfigured conditions, but can be overridden on the command line using the **encp** options described under section 3.2 *Encp Command Options* in 3.2.2 *Know what you're doing*. The numeric priority may change while the job waits in the queue depending on (a) the **encp** options used, (b) the elapsed time in the queue.

Users can manipulate the numerical priority number of their jobs using the **encp** options described under section 3.2 *Encp Command Options* in 3.2.2 *Know what you're doing*.

7.3 Fair Share Resource Allotment

There is a mechanism used in Enstore's queue management that prohibits any one storage group (experiment or group) from monopolizing tape drives. It is an algorithm we call "Fair Share". Fair Share determines which storage groups currently have jobs in progress (at a mover) and which ones don't, then gives preference to job requests associated with those that don't. This helps assure that there is a fair distribution of resources available to all groups currently using the system.

Under usual circumstances, storage groups may not reserve tape drives. However, in special cases, a storage group may request preallocation of tape drives.

7.4 Resource Ownership

When an experiment or group purchases one or more tape drives, the drives go into the pool of Enstore resources, accessible by all users, with the recognition that the purchasing group has the right to uninterrupted access to this number of drives at all times. The Fair Share configuration (section 7.3) gets modified to guarantee this access. The priorities of other jobs in the queue may be perturbed when Enstore needs to free up one or more tape drives for jobs submitted by the purchasing group.

In contrast to DAQ/Admin jobs, these (normal priority) jobs, identified by their storage group, must wait until a tape has dismounted in the normal way before being processed. A tape normally dismounts after all jobs in the queue requiring that tape have completed (this minimizes dead time associated with mounts and dismounts).